



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/721,271	11/26/2003	Kwang-cheol Oh	1033808-008	6479

21839 7590 10/17/2008
BUCHANAN, INGERSOLL & ROONEY PC
POST OFFICE BOX 1404
ALEXANDRIA, VA 22313-1404

EXAMINER

SHAH, PARAS D

ART UNIT	PAPER NUMBER
----------	--------------

2626

NOTIFICATION DATE	DELIVERY MODE
-------------------	---------------

10/17/2008

ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

ADIPFDD@bipc.com

Office Action Summary	Application No. 10/721,271	Applicant(s) OH ET AL.	
	Examiner PARAS SHAH	Art Unit 2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 08/01/2008.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-33 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-15 and 18-31 is/are rejected.
- 7) ☒ Claim(s) 16, 17, 32 and 33 is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413) |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date. _____ |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08) | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

1. This communication is in response to the Arguments filed on 08/01/2008. Claims 1-33 remain pending and have been examined. The Applicants' amendment and remarks have been carefully considered, but they are not persuasive and do not place the claims in condition for allowance.
2. All previous objections and rejections directed to the Applicant's disclosure and claims not discussed in this Office Action have been withdrawn by the Examiner.

Continued Examination Under 37 CFR 1.114

3. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 08/01/2008 has been entered.

Response to Arguments

3. Applicant's arguments filed on 08/01/2008 (pages 9-11) have been fully considered but they are moot in view of new grounds for rejection.

Response to Amendment

4. Applicants' amendments filed on 08/01/2008 have been fully considered. The newly amended limitations in claims 1, 5-7, 10, 12, 18, 22-25, and 29 necessitate new

Art Unit: 2626

ground for rejection. Hence, the prior art reference of Eryilmaz has been removed and the prior art reference by Nagasaki (US 6,629,070) has been applied.

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

6. Claims 1, 2, 5, 6, 18, 19, 22, and 23 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kushner *et al.* (US 6,321,197) in view of Durlach *et al.* (US 5,828,997) in view Nagasaki (US 6,629,070).

As to claims 1 and 18, Kushner *et al.* teaches a voice region detection apparatus, comprising:

a preprocessing unit for dividing an input voice signal into input frames (see col.4, lines 6-7, segments acquisition window into frames.) comprised of a sequence of elements having a number of runs (see Below mapping of Nagasaki);

a frame state determination unit for classifying the frames into voice frames and noise frames (see col. 4, lines 29-37, speech/noise classifier done by microprocessor 110) based on the random parameters extracted by the random parameter extraction unit; and

a voice region detection unit (see col. 5, lines 9-14, microprocessor 110) determines the starting point and ending point of the speech utterance.) for detecting a voice region by calculating start (see col. 6, lines 8-9 and Figure 2, microprocessor 110 determines the starting point and ending point of the speech utterance) and end positions of a voice based (see col. 6, lines 65-66 and Figure 2, endpoint is determined) on the voice and noise frames input from the frame state determination unit (e.g. From the determination of a speech utterance the voice regions are detected based on energy.).

However, Kushner *et al.* does not specifically teach the whitening unit for combining white noise to the input frames.

Durlach *et al.* does teach the whitening unit combining white noise to the input frames (see col. 5, lines 56-65 and Figure 2, target signals (speech) 50a, 50b, and 50n are added with the noise generator 60 by mixer 56). Although white noise is not used when adding to the target signals, it would have been obvious to add white noise to a signal or any other type of noise depending on environment simulated.

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the voice recognition as taught by Kushner *et al.* with the addition of a whitening unit as taught by Durlach *et al.* The motivation to have combined the references involve the ability to incorporate the directionality of a signal for sound localization (see Durlach *et al.*, col. 5, lines 60-

65) as would benefit the preprocessed signal from Kushner *et al.* for real-time environmental simulation.

However, Kushner *et al.* in view of Durlach *et al.* do not specifically teach the random parameter extraction unit.

Nagasaki does teach the random parameter extraction unit (see col. 6, lines 54-61 and col. 7, lines 1-5, thresholds used to determine whether voice is present or not present on the basis of energy) for extracting random parameters indicating the randomness of frames (see col. 6, lines 66-col. 7, lines 1-15 voice activity is detected based on the comparison to a threshold. The determination of energy is random since it is not known whether the frame is voice or noise. Further, the randomness is addressed by indicating the noise or voice present in the signal).

wherein the random parameter extraction unit extracts a random parameter for a frame input from the whitening unit (see col. 6, lines 54-60, voice presence determined based on intensity of energy in analysis region) based on a determination of the number of runs in said frame (see col. 7, lines 10-15, the number of runs is the energy values at each analysis region, where the frame is divided into (col. 4, lines 45-51))

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the voice recognition as taught by Kushner *et al.* in view of Durlach *et al.* with the addition of random parameter extraction unit as taught by Nagasaki for the purpose of accurately determining

voice frames and to prevent pulse noise to be considered as voice (see Nagasaki, col. 2, lines 22-36).

As to claim 2 and 19, Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki teach all of the limitations as in claims 1 and 18 above.

Furthermore, Kushner *et al.* teaches wherein the preprocessing unit samples the input voice signal according to a predetermined frequency (see col. 3, lines 66-col. 4, lines 10, digitize) and divides the sampled voice signal into a plurality of frames (see col. 4, lines 4-10, segmentation into frames is performed.) (e.g. The digitization of the voice signal makes the use of sampling frequency obvious as the signal is sent to the microprocessor for further processing. It is obvious that this sampling frequency is utilizing the Nyquist criterion.)

As to claim 4 and 21, Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki teach all of the limitations as in claims 1 and 18 above.

Furthermore, Durlach *et al.* teaches wherein the whitening unit comprises a white noise generation unit (see Figure 2, noise generator 60) for generating the white noise, and a signal synthesizing unit (see Figure 2, mixer 56) for combining the frames input from the preprocessing unit (see signals 50a, 50b, and 50n) with the white noise generated by the white noise generation unit (e.g. Noise is added to the target signal.).

As to claims 5 and 22, Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki teach all of the limitations as in claim 1 and 18 above.

Furthermore, Durlach *et al.* does teach the whitening unit combining white noise to the input frames (see col. 5, lines 56-65 and Figure 2, target signals (speech) 50a, 50b, and 50n are added with the noise generator 60 by mixer 56).

Furthermore, Nagasaki does teach each of said runs consists of consecutive identical elements in the sequence of elements that comprise the frame (see col. 7, lines 10-15, energy values are identically computed for each subframe).

As to claims 6 and 23, Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki teach all of the limitations as in claim 1 and 18 above.

Furthermore, Nagasaki does teach wherein the random parameter is : $NR=R/n$, where NR is a random parameter of a frame, n is a half of the length of the frame, and is the number of runs in the frame (see col. 7, lines 11-14, where the number of runs, specifically the cumulative sum is totaled (of four runs) and divided by the total number of sub-frames to determine the voice presence (4 sub-frames).).

7. Claims 3 and 20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki as applied to claims 2 and 19 above, and further in view of Mekuria (US 6,182,035).

As to claims 3 and 20, Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki teach all of the limitations as in claims 2 and 19 above.

However, Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki do not specifically teach the frames overlapping with one another.

Mekuria does teach the overlapping of frames (see col. 8, lines 28-29).

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have combined voice recognition as taught by Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki with the overlapping of frames as taught by Mekuria. The motivation to have combined the references involves the use of samples in more than one frame (see Mekuria col. 8, lines 28-29).

8. Claims 7-9 and 24-26 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki as applied to claims 1 and 18 above, and further in view of Pastor (US 5,572,623).

As to claims 7 and 24, Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki teach all of the limitations as in claims 1 and 18, above.

However, Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki do not specifically teach the voice frames including vocal frames and fricative frames.

Pastor does teach the frames including vocal and fricative frames (see col. 4, lines 66-67 and col. 5, lines 5-14).

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have combined the voice recognition as taught by Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki with the inclusion of fricative frames as taught by Pastor. The motivation to have combined the references involves the inclusion of fricatives that are present in at the start and end of speech (see Pastor col. 1, lines 29-33).

As to claims 8, 9, 25, and 26, Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki in view of Pastor teach all of the limitations as in claims 7 and 24, above.

Furthermore, Nagasaki teaches wherein the frame state determination unit (e.g. voice activity detector) determines if the random parameter of a frame extracted by is below a first threshold (see col. 7, lines 2-14, voice activity is detected based on the comparison to a threshold.) The determination of energy is random since it is not known whether the frame is voice or noise.) then it is a vocal frame (e.g. If the noise is below the value of the threshold, then speech is present or vocal frame. The use of a specific threshold would have been obvious to one skilled in the art in order to distinguish voice from noise. Hence, the use of below or above a threshold is matter of design choice and relativity. The Applicants do not indicate reasons for selecting the stated thresholds (see Applicant's Specification, page 11, lines 17-21).

9. Claims 10, 11, 27, and 28 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki in view of Pastor as applied to claims 8 and 25 above, and further in view of Chong-White *et al.* (US 7,065,485).

As to claims 10, 11, 27, and 28, Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki in view of Pastor teach all of the limitations as in claims 8 and 25, above.

However, Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki in view of Pastor do not specifically teach if the random parameter of a frame extracted by the random parameter extraction unit is above a second threshold, the relevant frame is a fricative frame.

Chong-White *et al.* does teach if the random parameter of a frame extracted by the random parameter extraction unit (see col. 7, lines 22-25, energy ratio computed, similar to Nagasaki) is above a second threshold (see col. 7, lines 46-47, fricatives identified when above a threshold), the relevant frame is a fricative frame. As to claims 11 and 28, it would have been obvious to select a threshold value for comparing different types of values of a signal with respect to a ratio.

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have combined the voice recognition as taught by Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki in view of Pastor with the inclusion of a threshold indicating a fricative as taught by Chong-White *et al.*

The motivation to have combined the references involves the ability to detect further unvoiced components in a signal consisting of speech and non-speech. Furthermore, the use of the voice recognition as taught by Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki in view of Pastor allows the ability to detect noise, voice and fricatives contained in the signal.

As to claims 12, 13, 29, and 30, Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki in view of Pastor in view of Chong-White teach all of the limitations as in claims 8 and 25, above.

Furthermore, Nagasaki teaches wherein the frame state determination unit determines that if the random parameter of the frame extracted by the random parameter extraction unit is below the second threshold, the relevant frame is a noise frame (see col. 7, lines 2-14, voice activity is detected based on the comparison to a threshold. The determination of energy is random since it is not known whether the frame is voice or noise).

However, Nagasaki does not specifically teach the use of two thresholds for comparison.

It would have been obvious to use multiple thresholds for classifying each frame so that the detection of voice and fricative frames can be detected as taught by Chong-White above in order to improve detection accuracy. Further, the values for the thresholds used are a matter of design choice based on the thresholds computed.

10. Claims 14 is rejected under 35 U.S.C. 103(a) as being unpatentable over Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki as applied to claim 1 above, and further in view of Rezayee *et al.* ("An Adaptive KLT Approach for Speech Enhancement").

As to claim 14, Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki teach all of the limitations as in claim 1, above.

However, Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki do not specifically teach a color noise elimination unit for eliminating color noise from voice.

Rezayee *et al.* teaches the enhancement of speech from colored noise (see Abstract).

It would have been obvious to one of ordinary skilled in the art to have combined the voice recognition as taught by Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki with the inclusion of a color noise eliminator as taught by Rezayee *et al.* the motivation to have combined the references is since colored noise consist of various noise variances and is not the same as white noise, which has same variance (see Rezayee *et al.* page 87, right column, 3rd paragraph, lines 12-17).

11. Claims 15 and 31 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki in view of Pastor in view

of Chong-White *et al.* (US 7,065,485) as applied to claims 10 and 27, above and further in view of Rezayee *et al.* ("An Adaptive KLT Approach for Speech Enhancement").

As to claims 15 and 31, Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki in view of Pastor in view of Chong-White *et al.* teach all of the limitations as in claim 1, above.

However, Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki in view of Pastor in view of Chong-White *et al.* do not specifically teach a color noise elimination unit for eliminating color noise from voice.

Rezayee *et al.* teaches the enhancement of speech from colored noise (see Abstract).

It would have been obvious to one of ordinary skilled in the art to have combined the voice recognition as taught by Kushner *et al.* in view of Durlach *et al.* in view of Nagasaki in view of Chong-White with the inclusion of a color noise eliminator as taught by Rezayee *et al.* the motivation to have combined the references is since colored noise consist of various noise variances and is not the same as white noise, which has same variance (see Rezayee *et al.* page 87, right column, 3rd paragraph, lines 12-17). Furthermore, it should be noted that the following elimination of colored noise is being done when speech is present. Hence, the detection of a vocal frame will entail speech is present and further enhance the signal from colored noise.

Allowable Subject Matter

12. Claims 16, 17, 32, and 33 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

13. The following is a statement of reasons for the indication of allowable subject matter: None of the prior art alone or in combination teaches the following limitations: $NR=R/n$, as recited in claims 6 and 23; “color noise ... obtained... amount of reduction in the random parameter...due to color noise” as recited in claims 16, 17, 32, and 33.

Conclusion

14. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Nakamura (US 5,937,375) is cited to disclose a voice presence activity detector based on power of subframes.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to PARAS SHAH whose telephone number is (571)270-1650. The examiner can normally be reached on MON.-THURS. 7:00a.m.-4:00p.m. EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Patrick Edouard can be reached on (571)272-7603. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Art Unit: 2626

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Paras Shah/
Examiner, Art Unit 2626

10/06/2008

/Patrick N. Edouard/
Supervisory Patent Examiner, Art Unit 2626